統計学Ｂ２

種村 秀紀　　千葉大学理学部数学・情報数理学科
〒 263-8522　千葉県千葉市稲毛区弥生町 1-33
http://www.math.s.chiba-u.ac.jp/ ~tanemura/index.html

平成 28 年 1 月 26 日

目 次

1　推定

統計学ではデータの解析を行う際, データは母集団と呼ばれる大きなデータ全体の一部分と考え, データを母集団から抽出された標本と呼ぶ. 母集団には大きく分けて有限母集団と無限母集団の二種類がある. 例えば N 人の総人口の中の n 人の身長データを抽出する場合は, N 人全体の身長からなるデータが有限母集団の例である. 一方, 同一条件で繰り返して実験や調査, 観測が行えると仮定できる場合に, ある確率分布からの無作為抽出により標本が得られると考えられ, この確率分布と, 得られるであろう無限の標本を同一視して無限母集団と呼ぶ.

得られた標本から母集団を特定する際に, 平均が未知の正規分布のように, パラメータ（母数） θ が未知であるような分布 P_θ が母集団であると仮定し, その θ をデータから特定することを推定（統計的推定）と呼び, そのための統計理論を推定理論と呼ぶ. また, パラメータ θ を「a 以上 b 以下」といった区間で推定するときを区間推定と呼び, 一意に推定するときをこれと区別するために点推定と呼ぶ.

（点）推定は標本の観測値 x が与えられたときにパラメータ θ のある推定値 $\hat{\theta}(x)$ を返すという意味で, 関数 $\hat{\theta}(\cdot)$ であると考えられる. ここで, この関数のことを推定量とよび, 推定値と区別することがあるが, 推定量と推定値という用語の使い分けはそれほど厳密にされていない.

1.1　十分推定量

表が出る確率が p, 裏が出る確率 $1-p$ の歪んだコインを n 回投げる試行, つまりパラメータ p のベルヌーイ分布からの n 回無作為抽出を考える. もし p が未知で, これを推定したいときには, n 回中に何回表が出たかが重要であり, その総回数さえわかれば何回目に表が出て何回目に裏が出たかは気にしないでよいということは直感的にわかる. このように, ある分布のパラメータ θ を推定したいときに, 分布から得られる標本 X のうち推定に十分な情報を含んだ統計量 $T=T(X)$ を十分統計量と呼び, 以下の式を満たす T として定義される.

$$P(X=x|T(X)=t,\theta)=P(X=x|T(X)=t). \tag{1.1}$$

$\boxed{\textbf{ex.}}$　i 番目のコインが表のときに X_i, 裏のときに $X_i=0$ とすると表の出た回数 $T(X)=\sum_{i=1}^n X_i$ は, 十分統計量である. 実際に条件付き確率を計算して, 確認してみよう. 条件づける前の確率を $\mu\in[0,1]$ と書くと

$$P(X=x|\mu)=\prod_{i=1}^n \mu^{x_i}(1-\mu)^{1-x_i}=\mu^{T(x)}(1-\mu)^{n-T(x)}=(1-\mu)^n\left(\frac{\mu}{1-\mu}\right)^{T(x)}$$

となる. よって $T(x)$ に関する確率と条件付き確率はそれぞれ

$$P(T(x)=t|\mu)={}_nC_t(1-\mu)^n\left(\frac{\mu}{1-\mu}\right)^{T(x)}$$

$$P(X=x|T(x)=t,\mu)=\frac{P(X=x,T(x)=t|\mu)}{P(T(x)=t|\mu)}=\frac{(1-\mu)^n\left(\frac{\mu}{1-\mu}\right)^t\mathbf{1}(T(x)=t)}{{}_nC_t(1-\mu)^n\left(\frac{\mu}{1-\mu}\right)^t}=\frac{1}{{}_nC_t}\mathbf{1}(T(x)=t)$$

となり, 上式の最右辺は確かにパラメータ μ によらないことが分かる. ただしここで, $\mathbf{1}(T(x)=t)$ は, $T(x)=t$ のとき 1 , それ以外のときは 0 の値をとる関数である. 十分統計量は一意ではないことは, 明らかである.

十分統計量は, 次の節で扱う尤度法による推定で発揮される. 逆にいうと, 十分統計量は 「尤度法を用いる限りにおいては」十分な統計量といえる. 実際, 尤度法の枠組みから外れた予測理論などでは十分統計量

以外の情報を用いると, さらに効率が上がる場合が知られている. しかし尤度法は実用上十分に汎用的で強力であるため, 十分統計量も十分に「十分な統計量」であるといえる.

単純なコイン投げの例で十分統計量かどうかの判断には計算を要したが, 簡単に判定する方法を与えてくれるのが以下に述べるフィッシャーーネイマンの分解定理である. 今, 標本 X の分布は密度関数 $f(x;\theta)$ をもつとする.

フィッシャーーネイマンの分解定理　$T(X)$ が θ の十分等計量であるとき, またそのときに限り

$$f(x;\theta) = h(x)g(T(x),\theta)$$

となる関数 h と g が存在する. つまり, 密度関数を θ によらない関数と, よる関数の積に分解したときに, 後者が $T(x)$ のみを含むような分解が存在する.

正規分布の例で実際に十分統計量を計算してみよう. $X_1,\ldots,X_n$ が正規分布 $N(\mu,v)$ から無作為抽出されているとすると, 密度関数は以下のようになる.

$$\begin{aligned} f(x;\mu,v) &= \frac{1}{(2\pi v)^{n/2}}\exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v}\right\} \\ &= \frac{1}{(2\pi v)^{n/2}}\exp\left\{-\frac{\sum_{i=1}^{n}x_i^2-2\mu\sum_{i=1}^{n}x_i+n\mu^2}{2v}\right\} \end{aligned}$$

まず μ の十分等計量を求めると

$$\frac{1}{(2\pi v)^{n/2}}\exp\left\{\frac{2\mu\sum_{i=1}^{n}x_i-n\mu^2}{2v}\right\}\times(\mu\text{ によらない部分})$$

という分解より, 分解定理を用いて $\sum_i X_i$ は μ の十分統計量であることがわかる.

一方, μ と v の両方をパラメータとするときには, ベクトル値をとる統計量

$$T(X) = (\sum_{i=1}^{n}X_i^2,\sum_{i=1}^{n}X_i)$$

が (μ,v) の十分統計量であることもいえる.

1.2　最尤推定量

確率密度関数を x が固定された θ の関数と考えたものを尤度関数とよび

$$L(\theta) = L(x,\theta) = f(x;\theta)$$

のように表す. ここで $L(x,\theta)$ は $f(x;\theta)$ と関数としては同じものだが, 統計学における重要性から, その用いられ方を含めて新たな名前がつけられている. また, 尤度関数 $L(x,\theta)$ を最大化するパラメータ $\hat{\theta}=\hat{\theta}(x)$ は x の関数となるが, これを最尤推定量と呼び, X が与えられたとき最尤推定値 $\hat{\theta}(X)$ が得られる.

ex.　正規分布 $N(\mu,v)$ から無作為抽出の標本 $X_1,X_2,\ldots,X_n$ がえられたとき, 平均 μ と分散 v の最尤推定量を求めてみよう. 密度関数

$$f(x;\mu,v) = \frac{1}{(2\pi v)^{n/2}}\exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v}\right)$$

の対数を μ で偏微分し,

$$\frac{\partial}{\partial\mu}\log f(X;\mu,v) = \frac{1}{v}\sum_{i=1}^{n}(X_i-\mu) = 0$$

より, $\hat{\mu}_n^{ML} = \overline{X}$ が得られ, μ の最尤推定量は標本平均となる．一方

$$\frac{\partial}{\partial v} \log f(X; \mu, v) = -\frac{n}{2v} + \frac{1}{2v^2} \sum_{i=1}^{n} (X_i - \mu)^2 = 0$$

を説くと $v = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2$ だが, $\hat{\mu}_n^{ML} = \overline{X}$ は分散 v によらなかったので, これを代入して,

$$\hat{v}_n^{ML} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

となり, 最尤推定量は（不偏ではない）標本分散と一致する．最尤推定量は, パレメータ変換による不変性をもつ．つまり, 分布をどのようなパラメータで表記しても, 最尤推定の結果で得られるぶんぷが同じになる．

1.3　不偏推定量

定義

推定量の期待値が推定しようとしている母数の値と一致しているときその推定量を 不偏推定量 という．

母数（パラメータ)

- 2 項分布 : p, n
- 正規分布 : μ(平均), σ(標準偏差)

母集団　$\mu = EX$: (母集団平均)

$\sigma = \sqrt{E[(X-\mu)^2]}$: (標準偏差)

標本　$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$: (標本平均)

$S_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2}$: (標本標準偏差)

不偏推定量の例

(1) 標本平均　$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ は μ の不偏推定量.

(2) 標本標準偏差 $S_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2}$ は σ^2 の不偏推定量.

証明 (1) 標本平均を繰り返してとると

$$\left.\begin{array}{lcl} \overline{X}_1 & = & \frac{1}{n} \sum_{i=1}^{n} X_i \\ \overline{X}_2 & = & \frac{1}{n} \sum_{i=1}^{n} X_{n+i} \\ \overline{X}_3 & = & \frac{1}{n} \sum_{i=1}^{n} X_{2n+i} \\ \vdots & & \\ \overline{X}_k & = & \frac{1}{n} \sum_{i=1}^{n} X_{(k-1)n+i} \end{array}\right\} \quad \text{これらの値も変動する.}$$

従って, 推定量 $\overline{X}$ も確率変数でありその平均を考えることができる. 計算をしてみると

$$E[\overline{X}] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$$

$$\therefore E[\overline{X}] = E[X]$$

(2) $\sigma^2 = E[(X - EX)^2] = E[X^2] - (EX)^2$

$$\begin{aligned}
&E[S_X^2] = \sigma^2 \text{を示す.}\\
&E[S_X^2] = E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2]\\
&\quad = \frac{1}{n-1}\sum_{i=1}^{n} E[X_i^2 + \overline{X}^2 - 2\overline{X}X_i]\\
&\quad = \frac{1}{n-1}\sum_{i=1}^{n} E[X_i^2] + \frac{1}{n-1}\sum_{i=1}^{n} E\left[\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2\right]\\
&\qquad - \frac{2}{n-1}\sum_{i=1}^{n} E\left[X_i \frac{1}{n}\sum_{j=1}^{n} X_j\right]\\
&\quad = \frac{n}{n-1}E[X_1^2] + \frac{n}{(n-1)n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} E[X_j X_k]\\
&\qquad - \frac{2}{(n-1)n}\sum_{i=1}^{n}\sum_{j=1}^{n} E[X_i X_j]\\
&\quad = \frac{n}{n-1}E[X_1^2] - \frac{1}{n(n-1)}\left\{\sum_{i=1}^{n} E[X_i^2] + \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} E[X_i X_j]\right\}\\
&\quad = \left(\frac{n}{n-1} - \frac{n}{n(n-1)}\right)E[X_1^2] - \frac{n(n-1)}{n(n-1)}E[X_1]^2\\
&\quad = \sigma^2
\end{aligned}$$

1.4 点推定と区間推定

$$\begin{array}{rcl}
\text{分布} & : & \text{平均 } \mu\text{, 標準偏差 } \sigma \text{ などの母数をもつ.}\\
\text{正規分布} & : & N(\mu, \sigma^2) \ \ \mu \text{ と } \sigma \text{ が母数 (パラメータ)}\\
\text{二項分布} & : & B_i(n, p) \ \ \left.\begin{array}{ll}\text{試行の回数} & n\\ \text{成功の確率} & p\end{array}\right\} \text{n と p が母数}
\end{array}$$

★ 母数が未知の時, n のおおよその値を <u>知りたい</u>. ⇒ (推定)

- **点推定** : 推定する母数の近似値
  - 平均 μ を推定するとき

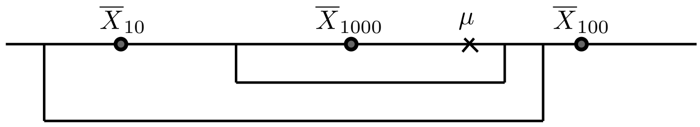
$$\text{標本平均} : \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

– 標本偏差 σ を推定するとき,

$$標本標準偏差 : S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

点推定は母数 θ を 1 つの値として推定するが, 区間推定は θ に対して確率の考え方を用いて推定を行う.

- **区間推定** : 予め定めた確率 (90%, 95%) で母数を含む区間 (**信頼区間**) を求める.



1.5 任意の分布での母集団平均 μ の推定

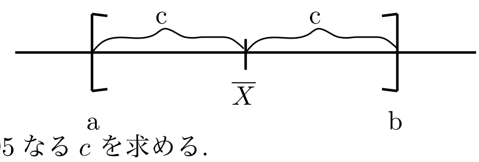
1.5.1 母集団分布の任意, 標準偏差 σ 既知

標本の数 $n \geq 25$ の時, 中心極限定理より標準偏差 $\overline{X}$ は平均 $\mu_{\overline{X}} = \mu$ (未知), 標準偏差 $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$ の正規分布で近似できる.

$\boxed{\textbf{目的}}$ 母数が含まれる確率が 95% である区間 $[a, b]$ を求める.

$a = \overline{X} - c$

$b = \overline{X} + c$



$P(\overline{X} - c \leq \mu \leq \overline{X} + c) \fallingdotseq 0.95$ なる c を求める.

$$\begin{aligned}
左辺 &= P(|\mu - \overline{X}| \leq c) \\
&= P\left(\left|\frac{\mu - \overline{X}}{\sigma_{\overline{X}}}\right| \leq \frac{c}{\sigma_{\overline{X}}}\right) \fallingdotseq P\left(|Z| \leq \frac{c}{\sigma_{\overline{X}}}\right)
\end{aligned}$$

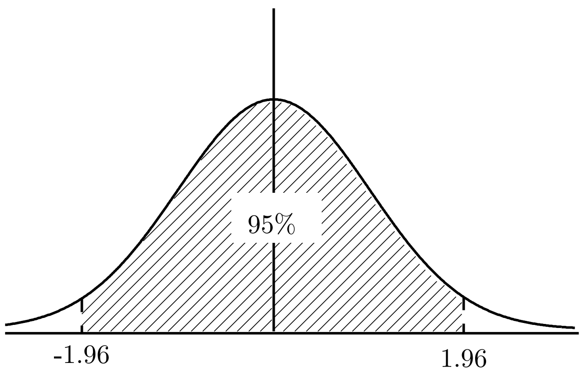
$\uparrow$ 中心極限定理の規準化

$$z = \frac{\mu - \overline{X}}{\sigma_{\overline{X}}}, \quad \left(\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{\pi}}\right)$$

従って $P(|Z| \leq x_{0.95}) \fallingdotseq 0.95$ となる $x_{0.95}$ を求めれば

$$c = \sigma_{\overline{X}} x_{0.95} = \frac{\sigma}{\sqrt{n}} x_{0.95}$$

となる. 標準正規分布表より $X_{0.95} = 1.96$.



μ が 95% で含まれる区間 (95% 信頼区間) は

$$\left[\overline{X} - 1.96 \times \frac{\sigma}{\sqrt{n}},\ \overline{X} + 1.96 \times \frac{\sigma}{\sqrt{n}}\right]$$

である．もし 90% 信頼区間を求めるときには $x = 1.645$ つまり

$$\left[\overline{X} - 1.645 \times \frac{\sigma}{\sqrt{n}},\ \overline{X} + 1.645 \times \frac{\sigma}{\sqrt{n}}\right]$$

とすればよい．

1.5.2 母集団分布が任意で標準偏差 σ が未知の場合

- $\sigma_{\overline{X}}$ の代わりに標準偏差 $S_{\overline{X}}$ を用いる．

$$\left(\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}},\ S_{\overline{X}} = \frac{S_X}{\sqrt{n}}\right)$$

- **近似定理**

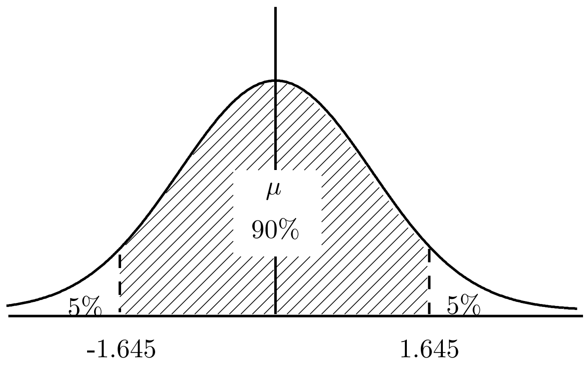
大きさ $n (n \geq 100)$ の標本をとると, 標本平均 $\overline{X}$ は平均 $\mu_{\overline{X}} = \mu$, 標準偏差 $S_{\overline{X}}$ の正規分布で近似できる．

(注) σ の点推定は S_X である．$n \geq 100$ の時, その誤差は無視できる．

μ の (近似的な)95% 信頼区間は

$$\overline{X} - 1.96 \times \frac{S_X}{\sqrt{n}} \leq \overline{X} + 1.96 \times \frac{S_X}{\sqrt{n}}$$

である．



ex. ある学校で 100 人の生徒の無作為標本が選ばれこれらの生徒に知能テストが行われた．テストの結果 100 人の生徒の知能指数が決まり $\overline{X} = 112, S_X = 11.0$ が得られた．これらの標本値を基にして, この学校の生徒全体の知能指数に対する 95% 信頼区間を求めよ．

$$\begin{aligned}
\overline{X} - 1.96 \times \frac{S_X}{\sqrt{n}} &\leq \mu \leq \overline{X} + 1.96 \times \frac{S_X}{\sqrt{n}} \\
112 - 1.96 \times \frac{11.0}{\sqrt{100}} &\leq \mu \leq 112 + 1.96 \times \frac{11.0}{\sqrt{100}} \\
109.844 &\leq \mu \leq 114.146 \\
109.8 &\leq \mu \leq 114.1
\end{aligned}$$

1.6　母集団の成功の割合 p の推定

1 回の成功の確率が p の試行を n 回繰り返す．標本からの成功の割合（標本支持率）を $\hat{p}$ と書く．つまり $\hat{p} = \frac{1}{n} \times$ 成功の回数.

成功の回数を S_n とおくと

$$P(S_n = k) = {}_n\mathrm{C}_k p^k (1-p)^{n-k}$$

つまり二項分布 $B(n,p)$ に従うことがわかる.

$$\begin{aligned} \mu_{S_n} &= E[S_n] &&= np \\ \sigma_{S_n} &= \sqrt{E[S_n^2] - \mu_{S_n}^2} &&= \sqrt{np(1-p)} \end{aligned}$$

$\hat{p} = \dfrac{S_n}{n}$ の平均は $\mu_{\hat{p}} = \dfrac{\mu_{S_n}}{n} = p$, 標準偏差は $\sigma_{\hat{p}} = \dfrac{\sigma_{S_n}}{n} = \sqrt{\dfrac{p(1-p)}{N}}$ となる.

前に述べたように, 標本数が大きいときは, 中心極限定理より標本平均の標準化 $\dfrac{\hat{p}-p}{\sigma_{\hat{p}}}$ は標準正規分布で近似できる．しかし, 母分散に当る $p(1-p)$ に未知のパラメータ p が入っていることに注意する必要がある. 近似的には p を $\hat{p}$ におきかえればよい.

- **近似定理**

  大きさ $n(n \leq 100)$ の標本をとり, 成功の割合を $\hat{p}$ とおく．$\hat{p}$ は平均 p, 標準偏差 $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ の正規分布で近似される.

母集団の成功の確率 p に対する 95%(90%) 信頼区間は

$$\hat{p} - \underset{(1.645)}{1.96} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + \underset{(1.645)}{1.96} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

となる.

$\boxed{\textbf{ex.}}$ ある都市で 1 日に少なくとも 1 箱のタバコを吸う成人男性の割合 p を推定したいと考えた．大きさ 300 の無作為標本を採って調べた結果, この様な喫煙者が 36 人いた．p の 95% 信頼区間を求めよ.

$$\begin{aligned} \hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &\leq p \leq \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ & \qquad\quad \left(\hat{p} = \frac{S_n}{n} = \frac{36}{300} = 0.12\right) \\ 0.12 - 1.96\sqrt{\frac{0.12 \cdot 0.88}{300}} &\leq p \leq 0.12 + 1.96\sqrt{\frac{0.12 \cdot 0.88}{300}} \\ 0.12 - 1.96\sqrt{0.000352} &\leq p \leq 0.12 + 1.96\sqrt{0.000352} \\ 0.083 &\leq p \leq 0.157 \end{aligned}$$

1.7　標本の大きさの問題

- 平均 μ の推定誤差 $|\mu - \overline{X}|$ がある値 ε を超えない確率を 95%(90%) 以上にしたい．標本の大きさ n をいくつ以上にすべきか.

– 95% の場合

$$\overline{X} - 1.96 \cdot \frac{S_x}{\sqrt{n}} \leq \mu \leq \overline{X} + 1.96 \cdot \frac{S_x}{\sqrt{n}}$$

つまり $|\mu - \overline{X}| \leq 1.96\dfrac{S_X}{\sqrt{n}}$ が 95% で起こる. 従って

$$\begin{array}{rcl} 1.96 \times \dfrac{S_X}{\sqrt{n}} & \leq & \varepsilon \\ & \Downarrow & \\ \left(\dfrac{1.96 S_X}{\varepsilon}\right)^2 & \leq & n \end{array}$$

とすればよい.

(※) 誤差を $\dfrac{1}{10}$ にしたい場合は標本の数は 100 倍にする必要がある.

- 割合 p の推定誤差 $|p - \hat{p}|$ がある値 ε を超えない確率を 95%(90%) 以上にしたい. 標本の数はいくつ以上にすべきか?

  – 95% の場合 $\hat{q} = 1 - \hat{p}$

$$\hat{p} - 1.96 \times \sqrt{\frac{\hat{p}\hat{q}}{n}} \leq p \leq \hat{p} + 1.96 \times \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$|p - \hat{p}| \leq 1.96 \times \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$1.96 \times \sqrt{\frac{\hat{p}\hat{q}}{n}} \leq \varepsilon$$

$$\left(\frac{1.96}{\varepsilon}\right)^2 \hat{p}\hat{q} \leq n$$

- $\hat{p}$ について何の情報も無い時, $\hat{p}$ のかわりに $\frac{1}{2}$ を使い,

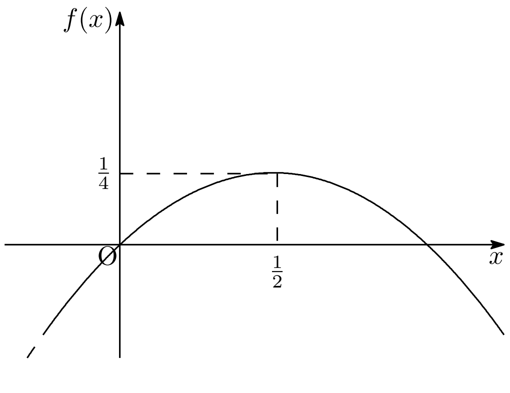
$$\frac{1}{4}\left(\frac{1.96}{\varepsilon}\right)^2 \leq n$$

とすればよい.

(理由) $\hat{p}\hat{q} = \hat{p}(1-\hat{p}) = f(\hat{p}), f(x) = x(1-x)$ の $0 \leq x \leq 1$ での最大値を求めると

$$f(x) = x(1-x) = -x^2 + x = -(x - \tfrac{1}{2})^2 + \tfrac{1}{4}$$

従って f は $x = \dfrac{1}{2}$ の時最大で最大値 $\dfrac{1}{4}$.



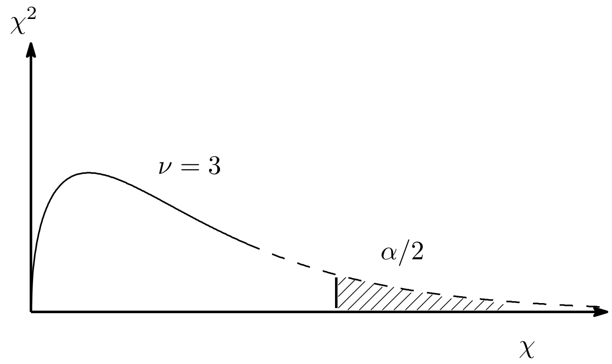
1.8　正規母集団の母平均, 母分散の区間推定

母分散の信頼区間

$Z_1, Z_2, \ldots, Z_k$ を独立な標準正規分布 $N(0,1)$ に従う確率変数とする.

$$\chi^2 = Z_1^2 + Z_2^2 + \cdots + Z_k^2$$

とすると, 確率変数 χ^2 が従う確率分布を 自由度 k の χ^2 分布($\chi^2(k)$ と表す) という. 自由度は独立な標準正規確率変数の数に対応している.



母集団分布が正規分布であるとき

$$\chi^2 = (n-1)S_X^2/\sigma^2$$

は自由度 $n-1$ の χ^2 分布 $\chi^2(n-1)$ に従う. χ^2 分布表からパーセント点 $\chi^2_{1-\alpha/2}(n-1)$, $\chi^2_{\alpha/2}(n-1)$ を求め

$$P(\chi^2_{1-\alpha/2}(n-1) \le (n-1)S_X^2/\sigma^2 \le \chi^2_{\alpha/2}(n-1)) = 1-\alpha$$

を得る. これを σ^2 について解くと母分散 σ^2 の信頼係数 $1-\alpha$ の信頼区間は

$$\left[\frac{(n-1)S_X^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S_X^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$$

となる.

母平均の信頼区間

母集団が正規分布に従うとき標本平均 $\overline{X}$ は正規分布 $N(\mu, \sigma^2/n)$ に従うからこれを標準化（基準化）すると

$$P(-Z_{\alpha/2} \le \tfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \le Z_{\alpha/2}) = 1-\alpha$$

$Z_{\alpha/2}$：標準正規分布表で得られる値

カッコ内の不等式を μ について解くと

$$(*) \qquad P(\overline{X} - \tfrac{Z_{\alpha/2}\sigma}{\sqrt{n}} \le \mu \le \overline{X} + \tfrac{Z_{\alpha/2}\sigma}{\sqrt{n}}) = 1-\alpha$$

となる. (標本標準偏差)^2=標準分散の標本分布を

$$S_X^2 = \tfrac{1}{n-1}(X_1-\overline{X})^2 + (X_2-\overline{X})^2 + \cdots + (X_n-\overline{X})^2$$

の平均 $E(S_X^2)=\sigma^2$ である事は示された（不偏性）

母平均 μ について調べるときに母分数, 母標準偏差がわかっているという状況は現実に想定しにくい. σ^2 の値が未知のままでは現実に用いる正確な標本分布を求める事ができない. したがって σ^2 の代わりに標本分散 S_X^2 を使う方法を考える.

スチューデントの t 統計量 (Student's t-statistic)

(a)　Z は標準正規分布 $N(0,1)$ に従う
(b)　Y は自由度 k の χ^2 分布 $\chi^2(k)$ に従う
(c)　Z と Y は独立である

とき

$$t=\frac{Z}{\sqrt{Y/k}}$$

と定義すると, t が従う確率分布を自由度 k の t 分布（スチューデントの t 分布）という.（$t(k)$ で表す）

k を大きくしていくと $t(k)$ は $N(0,1)$ に近くなる. (30 以上でかなり近く, ∞ で一致する)

自由度 k の t 分布 $t(k)$ の上側確率 $100\alpha\%$ のパーセント点を $t_\alpha(k)$ と書く.

スチューデントの t 統計量 (Student's t-statistic)

$$t=\frac{\overline{X}-\mu}{\sqrt{S_X^2/n}}$$

を定義する.

$$t=\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}\sqrt{S_X^2/\sigma^2}}=\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}\Big/\sqrt{\frac{\frac{(n-1)S_X^2}{\sigma^2}}{(n-1)}}$$

分子の分布は $N(0,1)$　$\underset{\text{独立}}{\Leftrightarrow}$　分母の二乗 $\times(n-1)$ の分布は $\chi^2(n-1)$

定理　母集団分布が正規分布で標準偏差 σ が未知のとき大きさ n の標本をとると

$$t=\frac{\overline{X}-\mu}{S_X}\sqrt{n}$$

は自由度 $\nu=n-1$ のスチューデント分布に従う.

μ の $100(1-\alpha)$ 信頼区間は

$$\overline{X}-t_\alpha(n-1)\frac{S_X}{\sqrt{n}}\leq\mu\leq\overline{X}+t_\alpha(n-1)\frac{S_X}{\sqrt{n}}$$

(σ が未知, 正規分布, 任意の n)

- μ の推定で n が小さく ($n<25$), σ が未知の時は大標本法（中心極限定理を用いたもの) を用いることができない. 一般にはこの場合は困難. もし "X が正規分布に従う" という条件があれば次の 小標本法 を用いる事ができる.

ex. ある製パン工場で作るパンの重さは正規分布に従うことがわかっている. 検査員がパンの重さを検査するために, 15 個の無作為標本をとった. これら 15 個のパンの重さの平均と標準偏差は 15.8 オンスと

0.3 オンスであった. μ の 95% 信頼区間を求めよ.

解答

$S = 0.3,\ \overline{X} = 15.8$
$t_{0.025}(14)$ の値を求める.

$$\left.\begin{array}{rcl} \nu = n-1 = 15-1 & = & 14 \\ p = (1-0.95)/2 & = & 0.025 \end{array}\right\} \longrightarrow t_{0.025}(14) = 2.145$$

従って

$$\begin{array}{rcl} 15.8 - 2.145 \times \dfrac{0.3}{\sqrt{15}} & \leq \mu \leq & 15.8 + 2.145 \times \dfrac{0.3}{\sqrt{15}} \\ \longrightarrow \quad 15.63 & \leq \mu \leq & 15.97 \end{array}$$

1.9 推定の演習問題

(i) ある地域の住民 400 人の無作為標本のうち 280 人は虫歯予防のため水道の水に少量のフッ素化合物を入れることを希望するという結果が出た. このデータからフッ素化合物の混合を希望する人の割合の 95% 信頼区間を求めよ.

(ii) 新聞社は消費税を支持する有権者の割合 p を推定しようと考えている. 0.90 の確率で推定値 $\widehat{p}$ を真の割合 p との差が 0.03 以内になるには標本を何人以上とらなければならないか？もし予備知識として $\widehat{p} \fallingdotseq 0.2$ を持っていればどれだけの標本でいいのか？

(iii) ある保険会社は長い間の経験から傷害保険加入者の 30% が 3 年間に少なくとも 1 回自動車事故を起こすことを知っている. この会社は, ある市の職員全員がこの傷害保険に加入する事を予期して, その場合の保険料率を決めたいと考えた. そこで, この市職員の中から 100 人の標本をとり, 過去 3 年間に少なくとも 1 回事故を起こした者を調べたら, 25 人がそうであった. 市職員は全保険加入者を代表すると仮定して, 次の問題を解け.

(a) この推定値の確率的な精度はいくらか. 推定値の誤差 ε を超えない確率が 95 パーセントとなるような ε の値を求めなさい. 実際の精度はいくらか. そして, このとき標本割合は理論から期待される結果に矛盾するかどうかを確かめなさい.

(b) 0.95 の確率で, 推定値の誤差を 0.03 以内にとどめるためには, 標本をさらに何人追加しなければならないか. p は未知であるとして解け.

(c) (b) の問題を, p の標本推定値を使う代わりに, 控えめな値 $p = 1/2$ を用いて解け.

(d) p の値は未知であるとして, p に対する 90% 信頼区間を求めよ. この区間は実際に p を含むか.

(iv) 1972 年秋発刊の医学雑誌 *Canadian Medical Association Journal* に, ”ビタミン C の大量投与が風邪に及ぼす効果について” という題の論文が掲載されている. この論文は, ビタミン C には風邪をはやく治す効果が若干あるという結論を下している. 風邪をひいて, ビタミン C の大量投与を受けた 407 人についての実験では, 風邪が治るまでに要した日数の平均と標準偏差はそれぞれ 5.25 日と 6.0 日であった. 標準偏差がこんなに大きいのは, おそらく, 風邪が治るまでに長い日数を要する人が必ず何人かいて, そのため風邪の回復に要する日数の分布は右側に長いすそを持つためである. このデータを用いて, 風邪が治るまでの日数の分布の平均値に対する 90% 信頼区間を求めよ.

(v) ある型の自動車の走行距離を推定するため, その型の車 30 台を標本に選び, 1 台ずつテストを行なった. 30 台の走行距離の平均と標準偏差がそれぞれ 19.6 マイルと 0.7 マイルになったとして, この型の車の平均走行距離に対する 90% 信頼区間を求めよ.

(vi) ある部品の生産者は, その製品には約 3% の不良品があると思っている. いま真の不良率を 0.97 の確率で 0.5% まで正確に推定したいとすれば, どれだけの標本をとらねばならないか.

(vii) 60 匹の実験動物を 2 週間, ある種の餌を与えて飼育した. そのときの体重増のデータから, $x = 42$ オンス, $s = 4$ オンスが得られた.

(a) 母集団平均の推定値として 42 オンスはどのくらい正確であるか. 推定値の誤差 ε を超えない確率が 95 パーセントとなるような ε の値を求めなさい.

(b) $\overline{x}$ と μ の違いを 0.95 の確率で 1/2 オンス以下にするためには, どれだけの標本が必要か.

(c) μ に対する 95% 信頼区間を求めよ. ここでは大標本法を用いよ.

(viii) 学生自治会は新しい学則を支持する学生の割合を推定しようとしている. 300 人の学生からなる無作為標本を選ぶことが提案された. これらの学生に対する質問の結果 $\hat{p} = 0.60$ を得た.

(a) この値は真の割合の推定値としてどのくらい正確といえるか. 推定値の誤差 ε を超えない確率が 95 パーセントとなるような ε の値を求めなさい .

(b) p の推定値を 0.04 以内の誤差で求めたいとすれば, 自治会は何人の標本学生をとればよいか. 0.95 の確率であれば十分であるとして, $\hat{p} = 0.60$ を用いて解け.

(c) 自治会が p の推定値として $\hat{p} = 0.60$ という予備知識を持っていなかったとすれば, どれだけの標本をとらねばならないか.

(ix) x が正規分布に従うとき, $x = 20, s = 4, n = 12$ が与えられたとして, スチューデントの t 分布を用いて

(a) μ に対する 95% 信頼区間を求めよ.

(b) μ に対する 99% 信頼区間を求めよ.

(x) ある銘柄のタバコ 20 本からなる標本について, そのニコチン含有量を調べ, $\overline{x} = 22(mg), s = 4(mg)$ を得た.

(a) スチューデントの t 分布を用いて, μ に対する 95% 信頼区間を求めよ.

(b) 大標本法によって解き, 大標本法と小標本法の結果を比較せよ.

(xi) ある種のガソリン添加物が自動車の走行距離を延ばすかどうかを調べるためのテストを行った. 25 台の車にそれぞれガソリン 5 ガロンを給油し, ガソリンが無くなるまで車を走らせた. 実験完了後, 各車ごとに 1 ガロンあたりの走行距離を計算した. こうして求めた 25 台の車の 1 ガロンあたり走行距離の平均と標準偏差はそれぞれ $\overline{x} = 18.5$ マイル, $s = 2.2$ マイルであった. 添加物を加えないで, 同じ種類の自動車を使い, 長期間にわたり行なってきたこれまでのテストの経験では, 1 ガロンあたり走行距離の平均と標準偏差は, 約 $\mu = 18.0$ マイル, $\sigma = 2.0$ マイルであった. この添加物は車の走行距離に影響しないと仮定して, 次の問題を解け.

(a) μ の推定値 $\overline{x}$ の確率的な精度を求めよ. また, $\overline{x}$ の実際の精度はどうなるか. 標本値は, 理論から期待される結果に矛盾しないか.

(b) 推定値の誤差が 1 ガロンあたり 1/2 マイルを超えない確率を 0.95 とするには, 何台の車でテストを行わねばならないか.

(c) μ に対する 95% 信頼区間を求めよ．求めた区間は実際に μ を含んでいるか.

(d) 添加物は平均, 分散のいずれにも影響しないという仮定をはずした上で, スチューデントの t 変数を使い, μ に対する 95% 信頼区間を求めよ.

(xii) 母分散 $\sigma^2 = 4$ の正規母集団から大きさ $n = 5$ の標本

$$9.75, \quad 7.95, \quad 12.80, \quad 8.25, \quad 9.86$$

を得た．母平均 μ の信頼係数 95 %の信頼区間を求めよ.

(xiii) 母分散 $\sigma^2 = 9$ の正規母集団から大きさ n の標本を抽出して, 母平均 μ の信頼係数 99 %の信頼区間を求めたい．その幅を 1 以下にするには, n をいくつ以上にすべきか？

(xiv) 母数 p の二項母集団 $Bi(1, p)$ から, $n = 50$ の大きさの標本を抽出して, 0 が 23, 1 が 27 であった．p の信頼係数 95 %の信頼区間を求めよ.

(xv) 1 時間ごとの受信電話数を記録したところ

$$4,\ 3,\ 5,\ 4.\ 8,\ 2,\ 5,\ 9,\ 3,\ 5$$

であった．ポアソン母集団 $Po(\lambda)$ を仮定して, λ の信頼係数 99 %の信頼区間を求めよ.

(xvi) 次のデータは, 20 匹のラットを 10 匹ずつ２群に分け, 一方には普通の食餌を与え, 他方には血中の赤血球を減らすと考えられる薬を混入した食餌を与えた場合の血液 $1mm^3$ 中の赤血球の数である.

投薬群と対照群のそれぞれの平均の差（すなわち, 薬の効果）の信頼係数 95% の信頼区間を求めよ.

投薬群	7.97	7.66	7.59	8.44	8.05	8.08	8.35	7.77	7.98	8.15
対照群	8.06	8.27	8.45	8.05	8.51	8.14	8.09	8.15	8.16	8.42

2　仮説の検定

ある確率変数 X ぼ分布が θ というパラメータによって規定されている場合を考える．このパラメータが属する集合（パレメータ空間と呼ぶ）Θ が互いに素な部分集合 (Θ_0, Θ_1) に分割されるとする．このとき，$H_0 : \theta \in \Theta_0$ という仮説が正しいかどうかをチェックしたいとする．この仮説を帰無仮説 (null hypothesis) といい，一方それとは別の仮説 $H_1 : \theta \in \Theta_1$ を対立仮説という．Θ_0 が１点 θ_0 からなる集合であるとき，すなわち $\Theta_0 = \{\theta_0\}$ であるとき，帰無仮説を単純帰無仮説であるという．単純対立仮説も同様に定義される．

帰無仮説の意味（百科事典マイペディア）：仮説検定で捨てるか捨てないか決めようとする仮説．差がない，効果がない，といった否定形のもの．帰無仮説が測定値によって捨てられれば，差がある，効果があるといった肯定的な結論が得られる．

特にパラメータが１次元，すなわち実数である場合は，帰無仮説 H_0 が $\theta = \theta_0$ の単純仮説に対して対立仮説が次の２つのいずれかであることが多い

$$H_1 : \theta \neq \theta_0 \tag{2.1}$$

$$H_1 : \theta > \theta_0 (\text{あるいは}\quad H_1 : \theta < \theta_0). \tag{2.2}$$

(2.1) 野場合の検定を両側検定，(2.2) の場合を片側検定とよぶ．

X の具体的なデータが得られたときに，これに基づいて帰無仮説が正しくないと判断することを，帰無仮説を棄却するといい，仮説が正しくないと結論づけるにはいたらない場合を採択（受容）するという．棄却にしろ採択にしろ，誤った判断をしてしまうことがあるが，帰無仮説が正しいにもかかわらず誤ってっこれを棄却してしまうことを **第一種の誤り**（第一種の過誤），対立仮説が正しいにもかかわれず誤って帰無仮説を採択してしまうことを**第二種の誤り**（第二種の過誤）という．

どちらの誤りが起こる確率も小さい方がよいが，通常第一種の誤りをおかす確率と第二種の誤りをおかす確率は相反の関係にあり，双方を同時に小さくすることはできない．そこで伝統的に，第一種の誤りをおかす確率を一定値 α 以下に抑えることを第一条件とし，その条件下でなるべく第二の誤りの確率を小さくするように工夫する．つまり，検定 δ について，真のパラメータが θ のときに帰無仮説を棄却する確率を $\beta_\delta(\theta)$ で表すとすると

$$\beta_\delta(\theta) \leq \alpha, \forall\theta \in \Theta_0$$

となる検定のみを考え，その中でなるべく良いものを選ぶことになる．この α のことを，**有意水準** とよぶ．有意水準として一般によく使われる値が決まっており，１％，５％，１０％の３つの数字が使われることが多い．第二種の誤りを小さくするということは，対立仮説が正しいときに帰無仮説を棄却する確率をなるべく大きくする，すなわち Θ_1 に属する θ についてなるべく $\beta_\delta(\theta)$ を大きくするということであるが，この確率を **検出力** とよび，θ の関数である $\beta_\delta(\theta)$ を検出力関数という．θ が実数の場合，θ の値を横軸にとって，検出力関数をグラフ化したものを検出力曲線とよぶ．

具体的な検定においては，X の値がとりうる空間（標本空間）は，２つに分割される．１つは採択域（受容域）と呼ばれる空間で，データがこの空間に値をとった場合は，帰無仮説は採択される．もう１つは棄却域と呼ばれる空間で，こちらにデータの値が属している場合は，帰無仮説は，棄却される．多くの場合，棄却域はある統計量，すなわち X の関数 $T(X)$ を使って

$$\{x|T(x) > C_1\}, \quad \{x|T(x) < C_1\}, \quad \{x|T(x) < C_1,\, T(x) > C_2\}$$

のような形で表されることが多い．このような場合 $T(X)$ を検定統計量，C_i $(I = 1, 2\ )$ を棄却限界という．c_i の具体的な値は，有意水準によって決まってくる．

$\boxed{\textbf{ex.}}$ 母集団は，$H_0 : Ge(1/6)$ または $H_1 : Ge(5/6)$ のどちらかであることは分かっている．これから２個の標本を取り出し，１個でも１以下なら帰無仮説 H_0 を棄却して対立仮説 H_1 を採択する．第一種の誤り α，第２種の誤り β，検出力 $1 - \beta$ をそれぞれ求めよ．また　棄却域を１個でも０なら H_0 を棄却するとしたときの α, β も求めよ．

解答　2つの標本を X_1, X_2 とすると, H_0 のもとでは X_1, X_2 は, パラメータ 1/6 の幾何分布に従う. よって

$$
\begin{aligned}
&\alpha = P(\{X_1 \le 1\} \cup \{X_2 \le 1\} | H_0) = P^{1/6}(\{X_1 \le 1\} \cup \{X_2 \le 1\}) \\
&= 1 - P^{1/6}(\{X_1 \ge 2\} \cap \{X_2 \ge 2\}) = 1 - P^{1/6}(\{X_1 \ge 2\}) P^{1/6}(\{X_2 \ge 2\}) \\
&= 1 - (5/6)^2 (5/6)^2 = \frac{671}{1296}
\end{aligned}
$$

H_1 のもとでは X_1, X_2 は, パラメータ 5/6 の幾何分布に従う. よって

$$
\begin{aligned}
&\beta = P(\{X_1 \ge 2\} \cap \{X_2 \ge 2\} | H_1) = P^{5/6}(\{X_1 \ge 2\} \cap \{X_2 \ge 2\}) \\
&= P^{5/6}(\{X_1 \ge 2\}) P^{5/6}(\{X_2 \ge 2\}) = (1/6)^2 (1/6)^2 = \frac{1}{1296}
\end{aligned}
$$

検出力は, $\frac{1295}{1296}$.

棄却域を 1 個でも 0 なら H_0 を棄却するとしたとき

$$
\begin{aligned}
&\alpha = P(\{X_1 \le 0\} \cup \{X_2 \le 0\} | H_0) = P^{1/6}(\{X_1 \le 0\} \cup \{X_2 \le 0\}) \\
&= 1 - P^{1/6}(\{X_1 \ge 1\} \cap \{X_2 \ge 1\}) = 1 - P^{1/6}(\{X_1 \ge 1\}) P^{1/6}(\{X_2 \ge 1\}) \\
&= 1 - (5/6)(5/6) = \frac{11}{36}
\end{aligned}
$$

であり

$$
\begin{aligned}
&\beta = P(\{X_0 \ge 1\} \cap \{X_2 \ge 1\} | H_1) = P^{5/6}(\{X_1 \ge 1\} \cap \{X_2 \ge 1\}) \\
&= P^{5/6}(\{X_1 \ge 1\}) P^{5/6}(\{X_2 \ge 1\}) = (1/6)(1/6) = \frac{1}{36}
\end{aligned}
$$

となる.

両側か片側か　一般に, 両側検定は, 母数 θ の値がある目標値 θ_0 と等しいかどうかを調べる場合に用いられる. 例えば, 工場で新しい機械を購入したとしよう. 機械が正しく働いていれば, 材料や運転条件によるばらつきがあるにしても, 製品は目標値の近くにあるはずである. 製品が目標値から大きく異なることは, 機械が正しく働いていないことを意味する. このような場合は, 両側検定となる.

片側検定は母数の大きさが理論的, 経済的に予測される場合に使われる. 例えば, 英語の特別授業の効果を調べる場合を考えよう. 英語の特別授業の前後での英語の試験の点数の平均を比べる場合, 特別授業に効果があれば試験後の点数が良くなっているはずである. このような場合, われわれが知りたいのは授業前後の得点が異なっているだけでなくて, 授業後の得点が向上したかどうかである. このような場合には, 対立仮説を不等号で与える片側検定を用いる.

2.1　平均の検定

$\boxed{\textbf{ex.}}$ 電球 A, B の平均寿命について

今まで使っている電球 A は平均 1180h, 標準偏差 90h の寿命を持つことがわかっている. セールスマンがやってきて「電球 A と同じ品質ですよ」と言って価格が安い電球 B を売りにきた. もし A と B が同じ品質であれば B を使いたい. B から大きさ $n = 100$ の標本平均 $\overline{X} = 1140h$ を得た. ただし B の電球の平均寿命の標準偏差は 90h とする.

【仮説 (Hypothesis)】

電球 A と電球 B は同じ寿命を持つ ($\mu_A = \mu_B$)

【対立仮説】

(i) 電球 A と電球 B の寿命は異なる. ($\mu_A \neq \mu_B$)　両側検定

$$\left.\begin{array}{ll}\text{(ii)} & \text{電球 A のほうが電球 B より寿命が長い. } (\mu_A > \mu_B) \\ \text{(iii)} & \text{電球 B のほうが電球 A より寿命が長い. } (\mu_B > \mu_A)\end{array}\right\} \text{片側検定}$$

今回の問題では事前の知識（セールスマンの話）より対立仮説として (ii) を用いる. 事前の知識が無い場合, 両側検定を用いる. なにかの事前の知識がある場合どちらかの片側検定を行う. したがって

$$\begin{array}{rcl}\text{仮説 } H_0 & : & \mu_B = 1180 \\ \text{対立仮説 } H_1 & : & \mu_B < 1180\end{array}$$

を $n = 100, \overline{X} = 1140$ に基づいて検定する.

【解法】

仮説 H_0 を認めた場合, 標本の結果が起こる確率がどのくらいかを計算して, あらかじめ与えた値より

$$\begin{array}{ccc}\text{大} & \longrightarrow & \text{採択} \\ \text{小} & \longrightarrow & \text{棄却}\end{array}$$

仮説を認めると電球 B の寿命 X の平均は 1180h, 標準偏差は 90h, この時, 中心極限定理を用いると $\overline{X}$ は平均 1180h, 標準偏差 $\dfrac{90}{\sqrt{100}} = 9h$ の正規分布で近似できる.　従って

$$P(\overline{X} \leq 1140) = P\left(\frac{\overline{X} - 1180}{9} \leq \frac{1140 - 1180}{9}\right)$$

$$P\left(Z \leq \frac{-40}{9}\right) \fallingdotseq P(Z \leq -4.44)$$

$$= P(Z \geq 4.44) \quad \text{これは非常に小さい値}$$

つまり仮説を仮定した場合, $\overline{X} = 1140$ 以下になる事は常識ではありえない.

- もし標本平均 $\overline{X} = 1160$ であれば

$$\begin{array}{rclcl}P(\overline{X} \leq 1160) & = & P(Z \leq \dfrac{-20}{9}) & & \\ & \fallingdotseq & P(Z \leq -2.22) & = & \text{0.5-0.4868} \\ & & & = & 0.0132 \\ & & & = & 1.32\% \end{array}$$

これでもかなり怪しい.
(セールスマンが嘘をついているらしい)

- もし標本平均 $\overline{X} = 1170$ であれば

$$\begin{array}{rcl}P(\overline{X} \leq 1170) & = & P(Z \leq -1.11) \\ & \fallingdotseq & \text{0.5-0.3665} \\ & = & 0.1335 \\ & = & 13.35\%\end{array}$$

このようなことはあり得る.

採択か棄却かを判定する基準 α を**有意水準**という.

α=0.05 の時

$$\left.\begin{array}{lll} \overline{X} = 1140 \text{ の時} & \longrightarrow & \text{棄却} \\ \overline{X} = 1160 \text{ の時} & \longrightarrow & \alpha > 0.013 \text{ より棄却} \end{array}\right\} \quad \text{A と B の寿命は異なると判定}$$

$$\overline{X} = 1170 \text{ の時} \longrightarrow \alpha < 0.1335 \text{ より採択} \qquad \text{A と B の寿命は異なると判断できない.}$$

棄却 : 有意である　　　採択 : 有意ではない

例のデータでは $\overline{X} = 1140$ であったので棄却. 従って B は A よりも劣っていると結論された.

α が与えられたとき,

$$\alpha = P\{\overline{X} < \overline{X_0} | H_0 \text{ が真 }\}$$

となる値 $\overline{X_0}$ が定まる. $\overline{X} < \overline{X_0}$ を棄却域, $\overline{X_0}$ を棄却域の境界という.

$$\alpha = 0.05 \quad P(Z \geq 1.645) = 0.05 \text{ より } Z_0(0.05) = 1.645$$

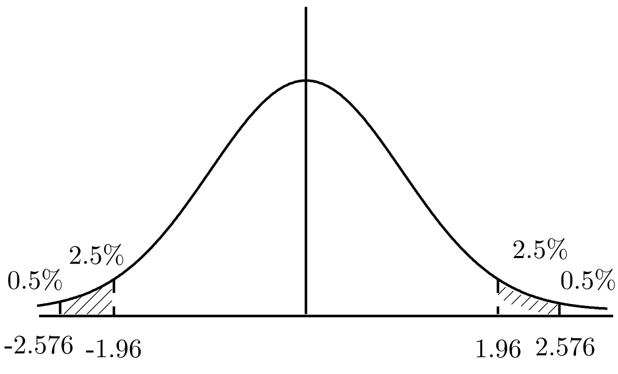
$$\begin{aligned} \frac{\overline{X}_0 - 1180}{9} &= -Z_0 \\ \overline{X}_0 &= 1180 \text{ - } 9 \times 1.645 \\ &= 1180 \text{ - } 14.8 \\ &= 1165.2 \end{aligned}$$

正規化した棄却域を用いると, 元の値に関係なく統一的に扱えるので便利である.

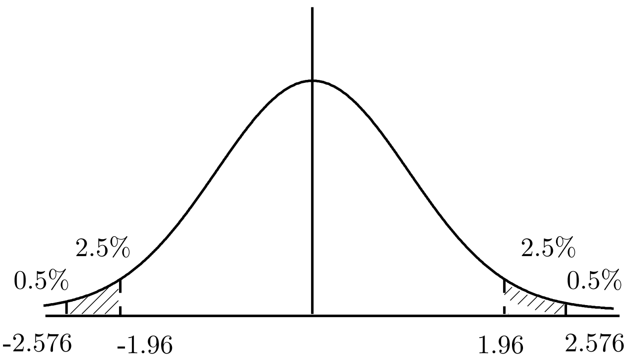
$$\begin{aligned} Z < Z_0 &\Rightarrow \text{棄却} \\ Z \geq Z_0 &\Rightarrow \text{採択} \end{aligned}$$

$\alpha = 1\%, 5\%, 10\%$ の場合の Z_0 の値を表にしてみると,

α	Z_0	
	両側検定	方側検定
0.01	2.576	2.326
0.05	1.96	1.645
0.10	1.645	1.286



両側検定



片側検定

2.2　平均の検定 (両側検定)

ex. ある学校で過去数年間の新入生についての記録では適性検査の得点は平均 115, 標準偏差 20 であった．今年の新入生に適性検査を行ったところ, 標本数 $n = 50$, $\overline{X} = 118$ を得た．今年の新入生が今までの学生と比べて優れているか劣っているかを検定せよ．(標準偏差は　と同じであるとしてよい.)

解答

$$\left.\begin{array}{ll} \text{仮説 } H_0: & \mu = 115(\text{例年と同じ}) \\ \text{対立仮説 } H_1: & \mu \neq 115(\text{異なる}) \end{array}\right\} \begin{array}{l} \text{を } n = 50, \overline{X} = 118 \text{ で検定する.} \\ (\sigma = 20 \text{ としてよい}) \end{array}$$

H_0 の下で $\overline{X}$ は平均 $\mu_{\overline{X}}$, 標準偏差 $\sigma_{\overline{X}} = \frac{20}{\sqrt{50}} = 2.83$ の正規分布で近似できる.

$\alpha = 0.05$ で両側検定すると, 正規化された棄却域は

$$|Z| > Z_0 = 1.960$$
$$|Z| = \left|\frac{118 - 115}{2.83}\right| = 1.06 < 1.960$$

従って H_0 を採択とした（有意ではない）　よって新入生は今までと異なると結論できない.

$$\text{有意水準 } \sigma = 0.1 \longrightarrow Z_0 = 1.645$$
$$\frac{\overline{X} - 115}{\sigma/\sqrt{n}} = \frac{118 - 115}{20/5\sqrt{2}} = \frac{3}{4}\sqrt{2}$$
$$Z_0 > \frac{\overline{X} - 115}{\sigma/\sqrt{n}} \longrightarrow \text{採択 (有意ではない)}$$

2.3　割合の検定

i 番目の試行を行ったとき

$$\begin{cases} \text{成功の場合} \longrightarrow & X_i = 1 \\ \text{失敗の場合} \longrightarrow & X_i = 0 \end{cases}$$

と X_i を定める.

n 回の試行のうち成功の回数 $= X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i$, 成功の確率 p, 失敗の確率を $q = 1 - p$ とすると, $Y = \sum_{i=1}^{n} X_i$ の分布は二項分布 $B_i(n,p)$ になる:

$$P(Y = k) = {}_n\mathrm{C}_k p^k q^{n-k}, \quad k = 0, 1, \ldots, n$$

二項分布 $B_i(n,p)$ の平均は np, 標準偏差は $\sqrt{npq}$ であるから

$$\hat{p} \equiv \quad \text{成功の割合} \equiv \frac{\text{成功の回数}}{n}$$

の平均は p, 標準偏差は $\sqrt{\frac{pq}{n}}$ であることが分かる.

ex. 飲酒運転率 (両側検定)

1974 年 7 月 11 日付の Los Angeles Times には, カリフォルニア大学医療心理学教授のアービン氏がオレンジ郡保健局と協力して, オレンジ郡の週末の運転者中の飲酒運転者の数を調べた調査報告が掲載されている. アメリカ全国の飲酒運転率は約 5% であるという. オレンジ郡で停車を命ぜられて検査を受けた 1000 人のドライバーのうちの 7% が飲酒運転であると判定された. 飲酒運転かどうかの判定は血液中のアルコールの量によってなされる. オレンジ郡のドライバーの真の飲酒運転率は全国のドライバーのそれと同じであるという仮説を, 同じでないという対立仮説に対して, $\alpha = 0.05$ で検定せよ.

解答

カリフォルニア州の飲酒運転率 p とおく.

$$\left.\begin{array}{ll} \text{仮説 } H_0 & : p = 0.05 \\ \text{対立仮説 } H_1 & : p \neq 0.05 \end{array}\right\} \text{として仮説検定する.}$$

仮説 H_0 の下で n は十分大きいので
$\hat{p}$ は平均 0.05, 標準偏差 $\sqrt{\frac{0.05 \times 0.95}{1000}} \fallingdotseq 0.0065$, である正規分布で近似される. $\alpha = 0.05$, 両側検定での棄却域は $|Z| > 1.960$

$$|Z| = \frac{0.07 - 0.05}{0.0063} = 2.9$$

従って

$$|Z| > Z_0 = 1.960 \longrightarrow H_0\text{は棄却}$$

よってカリフォルニア州の飲酒運転率は全州と異なると結論された.

2.4 2 つの平均値差の検定

2 つの母集団があり,

$$\begin{cases} \text{母集団 } 1: & \text{平均}\mu_1 \quad \text{標準偏差}\sigma_1 \\ \text{母集団 } 2: & \text{平均}\mu_2 \quad \text{標準偏差}\sigma_2 \end{cases}$$

であるとする.

母集団 i から大きさ n_i の標本をとり, 標本平均 $\overline{X}_i$, 標本標準偏差 S_i を得た. $(i = 1, 2.)$

- **近似定理**

  母集団分布が共に任意で母集団標準偏差 σ_1, σ_2 が未知のとき, 大きさ $n_1, n_2 (n_1, n_2 \geq 50)$ の標本をとると, 2 つの平均の差 $\overline{X}_1 - \overline{X}_2$ の分布は平均 $\mu_1 - \mu_2$, 標準偏差 $\sqrt{S_1^2/n_1 + S_2^2/n_2}$ である正規分布で近似される.

ex. 市役所の電球の購入

市役所は 2 つの銘柄の電球⊠, ⊠のどちらかを購入する事にした. そこで電球⊠, ⊠の平均寿命を調べる.
電球⊠の $n_1 = 100$ の標本から $\overline{X}_1 = 1160h, S_1 = 90h$
電球⊠の $n_2 = 100$ の標本から $\overline{X}_2 = 1140h, S_1 = 80h$
を得た. 2 つの銘柄に差があるかどうか

両側検定を行う.

電球⊠, ⊠の平均分布をそれぞれ μ_1, μ_2 とおく. そして

$$\begin{cases} \text{仮説 } H_0 & \mu_1 = \mu_2 \quad \text{(2 つの銘柄に品質の差は無い.)} \\ \text{対立仮説 } H_1 & \mu_1 \neq \mu_2 \quad \text{(2 つの銘柄の品質に差がある.)} \end{cases}$$

とする.

H_0 の下で $\overline{X}_1 - \overline{X}_2$ の分布は平均 0, 標準偏差 $\sqrt{\dfrac{90^2}{100} + \dfrac{80^2}{100}} \fallingdotseq 12$ の正規分布で近似できる.

- $\alpha = 0.05$ で両側検定する.

$$\text{棄却域は } |Z| > Z_0 = 1.96$$
$$\left| \frac{1160 - 1140}{12} \right| \fallingdotseq 1.67 < 1.96$$

  採択となる. 従って "2 つの銘柄に差があるとはいえない" と結論される.

- $\alpha = 0.10$ で両側検定する.

$$\text{棄却域は } |Z| > Z_0 = 1.645$$
$$\left| \frac{1160 - 1140}{12} \right| \fallingdotseq 1.67 > 1.645$$

  棄却となる. 従って "2 つの銘柄には差がある" と結論される.

2.5 2 つの割合の差の検定

母集団 i では母集団の成功の割合が p_i である. ここから n_i 個の標本をとる. この時成功の回数を Y_i, 成功の割合を $\hat{p}_i \left(= \dfrac{Y_i}{n_i} \right)$ と書く. $(i = 1, 2.)$

- **近似定理**

  母集団 1 と 2 の成功の割合 p_1 と p_2 が共に未知のとき大きさ $n_1, n_2 (n_i \geq 50, n_2 \geq 50)$ の標本をとり, 標本からの成功の割合 $\hat{p}_1, \hat{p}_2$ を得た. 確率変数 $\hat{p}_1 - \hat{p}_2$ は平均 $p_1 - p_2$, 標準偏差 $\sqrt{\dfrac{\hat{p}_1 \hat{q}_1}{n_1} + \dfrac{\hat{p}_2 \hat{q}_2}{n_2}}$ の正規分布で近似できる. ただし, $\hat{q}_i = 1 - \hat{p}_i$.

  特に仮説 $H_0 : p_1 = p_2 = p$ のもとでは

$$\begin{cases} \text{平均：} & p_1 - p_2 & = p - p = 0 \\ \text{標準偏差：} & \sqrt{\dfrac{\hat{p}\hat{q}}{n_1} + \dfrac{\hat{p}\hat{q}}{n_2}} & = \sqrt{\dfrac{(n_1+n_2)\hat{p}\hat{q}}{n_1 n_2}} \end{cases}$$

ここで $\hat{p} = \dfrac{Y_1 + Y_2}{n_1 + n_2}, \hat{q} = 1 - \hat{p}$ とできる.

$\boxed{\textbf{ex.}}$ ビタミン C の風邪に及ぼす効果 (片側検定)

ビタミン C の大量投与を受けた 407 人のうち, 期間中 105 人が風邪にかからず, 偽薬投与を受けた 411 人のうち 76 人が風邪にかからなかった. ビタミン C の風邪の予防効果を調べよ.

$$\begin{cases} p_1 & : & \text{ビタミン C 投与を受けた集団が風邪にかからない割合} \\ p_2 & : & \text{偽薬投与を受けた集団が風邪にかからない割合} \\ \text{仮説 } H_0 & : & p_1 = p_2 (\text{ビタミン C は風邪に効果なし}) \end{cases}$$

標本は $\left(\begin{array}{lll} n_1 = 407, & Y_1 = 105, & \hat{p}_1 = \dfrac{105}{407} \fallingdotseq 0.26 \\ n_2 = 411, & Y_2 = 76, & \hat{p}_2 = \dfrac{76}{411} \fallingdotseq 0.18 \end{array} \right.$

$$\hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2} = \frac{181}{818} \fallingdotseq 0.22$$

$\hat{p}_1 - \hat{p}_2$ の分布は

$$\text{平均 } 0, \quad \text{標準偏差} \sqrt{\frac{0.22 \times 0.78}{407} + \frac{0.22 \times 0.78}{411}} = 0.029$$

の正規分布で近似される.

- $\alpha = 0.05$ で片側検定する.

  棄却域は $|Z| > Z_0 = 1.645$

$$Z = \left| \frac{0.26 - 0.18}{0.029} \right| \fallingdotseq 2.76 > 1.645$$

  ゆえに H_0 は棄却. 従って "ビタミン C は風邪の予防に効果がある" と結論された.

2.6 正規母集団に対する仮説検定

2.6.1 母分散に対する仮説検定

母分散 σ^2 に対する帰無仮説 $H_0 : \sigma^2 = \sigma_0^2$ の検定は, 標本分散 s^2 によって, 検定統計量

$$\chi^2 = (n-1)s^2/\sigma_0^2 = \frac{\sum_{j=1}^{N} (\overline{X} - X_j)^2}{\sigma_0^2}$$

が, 帰無仮説のもとで自由度 $(n-1)$ の χ^2 分布 $\chi^2(n-1)$ に従うという性質を用いて行われる. 検定の有意水準を α とするとき,χ^2 分布表から $\chi^2(n-1)$ のパーセント点 $\chi^2_{1-\alpha}$, χ^2_α などを求めておき,

(a) 対立仮説が $H_1 : \sigma^2 \neq \sigma_0^2$ のときは, （両側検定）

$\chi^2_{1-\alpha}(n-1) < \chi^2 < \chi^2_{\alpha}(n-1)$ のときは, H_0 を棄却せず, それ以外は棄却する.

(b) 対立仮説が $H_1 : \sigma^2 > \sigma_0^2$ のときは, （右片側検定）

$\chi^2_{\alpha}(n-1) < \chi^2$ のときは, H_0 を棄却し, それ以外は棄却しない.

(c) 対立仮説が $H_1 : \sigma^2 < \sigma_0^2$ のときは, （左片側検定）

$\chi^2 < \chi^2_{1-\alpha}(n-1)$ のときは, H_0 を棄却し, それ以外は棄却しない.

これらの検定を正規母集団の母分散についての χ^2 **検定** (**χ^2-test**) という.

$\boxed{\textbf{ex.}}$ (能力のばらつきの検定)

ある小学校では入学時に知能テストを行っていたが, 従来は平均 50 で分散 36 であった. 本年度の児童について 25 人をランダムに選び, 例年と同じ条件でテストをしたところ. 平均が 53 で分散 48 を得た. 本年度は児童の揃い方が例年と違うと見てよいか？

帰無仮説 $H_0 : \sigma^2 = 36$ を対立仮説 $H_1 : \sigma^2 \neq 36$ に対して, 有意水準 10 % で検定する. $\chi^2 = \frac{24\times 48}{36} = 32$ であるが, $\chi^2_{0.95}(24) = 13.848$, $\chi^2_{0.05}(24) = 36.415$ であるから, 有意水準 10 % で, H_0 は棄却されない. 本年度の児童の質の揃い方が特に例年とかわっているとはいえない.

なお, 分散が大きいことが特別の対応を要するときは, 対立仮説を $H_1 : \sigma^2 > 36$ とすればよい.

2.6.2 母平均の差の検定

２つの正規母集団の母平均の差の検定は, 実用上にも非常に重要である. 例えば, 新しい治療法の効果を調べる場合を考え, 患者を２つのグループに分けて, 一方のみに新しい治療を行い, ２つのグループで結果に差があるかどうかを検定する. これを２標本検定 (two-sample test) という. 治療法を行ったグループを一般に処理群 (treatment group), 行わずに比較対照の基準としておき, 実験の管理するためのグループを対照群 (contrast group) あるいは制御群 (control group) という.

２つの正規母集団 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ をそれぞれの大きさ m, n の標本 $X_1, X_2, \ldots, X_m$, $Y_1, Y_2, \ldots, Y_n$ を抽出したとしよう. このとき, 帰無仮説は

$$H_0 : \mu_1 = \mu_2$$

であり, 対立仮説は

(a) 両側ならば $H_1 : \mu_1 \neq \mu_2$,

(b) 片側ならば $H_1 : \mu_1 > \mu_2$ または $H_1 : \mu_1 < \mu_2$

である. 対立仮説が両側か片側かは, 他の検定と同様その目的から決定される.

検定方法は, ２つの分散が等しいかどうかによって, (I), (II) の２つに分けられる.

(I)　２つの分散が等しく, $\sigma_1^2 = \sigma_2^2 = \sigma^2$ であるときは, ２つの標本平均を $\overline{X}, \overline{Y}$ とし, σ^2 を

$$s^2 = \frac{\sum_{i=1}^{m}(X_i - \overline{X})^2 + \sum_{j=1}^{n}(Y_j - \overline{Y})^2}{m+n-2} = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2}$$

で推定すると, H_0 のもとでは２標本 t 統計量

$$t = \frac{\overline{X} - \overline{Y}}{s\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

は, 自由度 $m+n-2$ の t 分布 $t(m+n-2)$ に従う. よって (a) では両側検定となり, $|t| > t_{\alpha/2}(m+n-2)$ のとき帰無仮説を棄却し, それ以外は棄却しない. また, (b) では片側検定となり, それぞれ, $t > t_{\alpha}(m+n-2)$ のとき帰無仮説を棄却し, それ以外は棄却せず（右片側）, $t < -t_{\alpha}(m+n-2)$ のとき帰無仮説を棄却し, それ以外は棄却しない（左片側）.

$\boxed{\textbf{ex.}}$ (対照実験)

次のデータは, 20 匹のラットを 10 匹ずつ２群に分け, 一方には普通の食餌を与え, 他方には血中の赤血球を減らすと考えられる薬を混入した食餌を与えた場合の血液 $1mm^3$ 中の赤血球の数である.

投薬群	7.97	7.66	7.59	8.44	8.05	8.08	8.35	7.77	7.98	8.15
対照群	8.06	8.27	8.45	8.05	8.51	8.14	8.09	8.15	8.16	8.42

$m = n = 10$, $\overline{X} = 8.004$, $\overline{Y} = 8.230$, $\overline{X} - \overline{Y} = -0.226$. また, $s_1^2 = 0.0761, s_2^2 = 0.0294, s^2 = 0.0528$, $s = 0.230$. よって $t = -2.200$. $t_{0.025}(18) = 2.101$, $t_{0.05}(18) = 1.734$ から, 帰無仮説 $H_0 : \mu_1 = \mu_2$ は, 両側検定（有意水準５％）でも左片側検定（有意水準５％）でも棄却され, 薬効はあると認められる.

(II) 両母集団の母分散が等しくない場合は少し複雑になる.

$$s_1^2 = \sum \frac{(X_i - \overline{X})^2}{m-1}, \quad s_2^2 = \sum \frac{(Y_i - \overline{Y})^2}{n-1}$$

とすると, 帰無仮説が正しい場合

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

は近似的に自由度が

$$\nu = \frac{(s_1^2/m + s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}}$$

に最も近い整数 $\nu^\star$ の自由度の t 分布 $t(\nu^\star)$ に従う. したがって t 値をデータから計算すれば,

(a) では両側検定となり, $|t| > t_{\alpha/2}(\nu^\star)$ のとき帰無仮説を棄却し, それ以外は棄却しない.

(b) では片側検定となり, それぞれ, $t > t_\alpha(\nu^\star)$ のとき帰無仮説を棄却し, それ以外は棄却せず（右片側）, $t < -t_\alpha(\nu^\star)$ のとき帰無仮説を棄却し, それ以外は棄却しない（左片側）.

この検定は, **ウェルチの検定 (Welch's test)** と呼ばれる.

2.6.3 母分散の比の検定

２つの正規母集団の母平均が等しいか否かの検定方法は, 母分散が等しい（母分散の比が１）かどうかによって異なるため, 分散が等しいかどうかの検定が必要である. また, 母分散が等しいことそれ自体に意味があることもある. （小学校の２クラスの能力のばらつきが等しいか否かなど）. この帰無仮説は

$$H_0 : \sigma_1^2 = \sigma_2^2$$

対立仮説は

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

である. ここで, H_0 のもとでのフィッシャーの分散比を

$$F = \frac{s_1^2}{s_2^2}$$

ただし

$$s_1^2 = \sum_{i=1}^{m} \frac{(X_i - \overline{X})^2}{m-1}, \quad s_2^2 = \sum_{i=1}^{n} \frac{(Y_i - \overline{Y})^2}{n-1}$$

とすると, 帰無仮説が正しい場合, F は自由度 $(m-1, n-1)$ の F 分布 $F(m-1, n-1)$ に従う. したがって, この F の値が

$$F_{1-\alpha/2}(m-1, n-1) \leqq F \leqq F_{\alpha/2}(m-1, n-1)$$

であるときは帰無仮説を棄却せず,

$$F < F_{1-\alpha/2}(m-1, n-1) \quad \text{または} \quad F > F_{\alpha/2}(m-1, n-1)$$

であるときは帰無仮説を棄却する.

これら, F 統計量を検定統計量として用いる検定を一般に **F 検定** という. なお, 片側対立仮説

$$H_1 : \sigma_1^2 > \sigma_2^2$$

なども考えられ, 片側 F 検定という.

ex. 対照実験で, $F = 0.0761/0.0294 = 2.59 < F_{0.025}(9, 9) = 4.026$. これから, 有意水準 5 %で, 両母分散は等しいという仮説は採択された.

2.7 χ^2 検定

χ^2 分布は母分散についてのいろいろな検定に用いられたが, 広く"ばらつき" についての検定の基準としても, 近似的に用いられる. そのうち, 確率分布の適合度の検定および独立性の検定はとくに重要なものである.

2.7.1 適合度検定

仮定された理論上の確率分布に対して, 標本から求めれた度数が適合するか否かを検証するのが, **適合度の χ^2 検定**（χ^2-test of goodness of fit ）である.

標本の大きさが n, 各カテゴリーに属する確率が $p_i, i = 1, 2, \ldots, k$ である多項分布を考える. n 個の観測値中, 各カテゴリーに属する個数を $X_i, i = 1, 2, \ldots, k$ とする. これをもとに検定を行う. 最初に一般的な帰無仮説について考える. θ を h 次元のパラメータとして, この θ によって, 各 p_i が表されている状態を帰無仮説とし, 対立仮説は帰無仮説以外の状態とする. すなわち

$$H_0 : p_i = p_i(\theta) \quad \text{v.s.} \quad H_1 : H_0 \text{ ではない}$$

という仮説検定問題を考える. このとき, 帰無仮説のもとでの θ の最尤推定量を $\hat{\theta}$ とし, 各項目の期待度数を $\hat{x}_i = n\hat{p}_i(\hat{\theta})$ として定義する. このとき検定統計量 W を

$$W = \sum_{i=1}^{k} \frac{(x_i - \hat{x}_i)^2}{\hat{x}_i}$$

として定義する. 帰無仮説が正しければ, 期待度数 $\hat{x}_i$ と標本での値（観測度数とよぶ）x_i の差は小さくなるはずであり, 全体として W は小さくなると考えられることから, W が大きいときに棄却する検定が自然である. 有意水準にあわせて棄却限界を決める必要があるが, 実は n が大きいときに W は自由度 $k-h-1$ のカイ二乗分布で近似できることが知られているので, これを利用して

$$W > \chi_\alpha^2(k-h-1)$$

を棄却域にすれば, n が大きいとき, ほぼ有意水準 α の検定となる. この検定のことを（カイ二乗）適合度検定と呼ぶ.

次の具体例を考えてみる.

ある商品の購買客は 20 代が 30 代の２倍であるといわれている．このことを確かめるために一定期間調査を行ったところ, 20 代の客が 52 人, 30 代の客が 30 人, それ以外の年代の客が 15 人となった．この結果のもとに「20 代の客が 30 代の客の２倍いる」という仮説を統計的に検定するとどうなるだろうか？もし, 調査の規模が 10 倍でそれぞれの数が 520 人, 300 人, 150 人だったらどうなるか．２つの場合で結論に差が出るとしたら, どう説明すればよいだろうか．さらに条件を増やした仮説「20 代の客, 30 代の客, それ以外の年代の客を割合が, 2:1:1 である」の場合は, 結論はどうなるだろうか.

この適用場面を使って有意水準５％の検定を行ってみる.

$$p_1 = 20\text{ 代の客の割合},\ p_2 = 30\text{ 代の客の割合},\ p_3 = \text{それ以外の年代の客の割合}$$

とする．この場合, 帰無仮説は「20 代の客が 30 代の客の２倍」は, １次元のバラメータ θ を使うと

$$H_0 : p_1 = 2\theta,\ p_2 = \theta,\ p_3 = 1 - 3\theta$$

となる．この帰無仮説のもとで, 多項分布の尤度

$$p(x_1, x_2, x_3 : \theta) = \frac{n!}{x_1!x_2!x_3!}(2\theta)^{x_1}(\theta)^{x_2}(1-3\theta)^{x_3}$$

を最大にする θ の値（最尤推定量）は,

$$\hat{\theta} = \frac{x_1 + x_2}{3(x_1 + x_2 + x_3} = \frac{82}{97 \times 3}$$

となる．これを代入すると, 期待度数は,

$$\hat{x_1} = 2 \times \frac{82}{97 \times 3} = \frac{164}{3},\ \hat{x_2} = \frac{82}{97 \times 3} \times 97 = \frac{82}{3},\ \hat{x_3} = 97 - \hat{x_1} - \hat{x_2} = \frac{45}{3}$$

となり, 検定量の値は,

$$W = \frac{(52 - 164/3)^2}{164/3} + \frac{(30 - 82/3)^2}{82/3} + \frac{(15 - 45/3)^2}{45/3} = \frac{16}{41} \sim 0.390$$

である．これを自由度 $k - h - 1 = 3 - 1 - 1 = 1$ のカイ二乗分布の上側５％点 3.841 と比較すると小さいので, 帰無仮説は採択される．（ちなみに, １０％検定を行っても棄却限界は 2.706 なので, 仮説は採択される.）

標本サイズが比率は同じまま 10 倍になった場合は, W の値は 10 倍の $16041 \sim 3.90$ になることはすぐ分かる．これは, 3.841 より大きいので, 帰無仮説は棄却されることになる．２つの場合で結論が分かれたが, これは同じ割合の標本であってもその意味するところが違うためである．調べた客の数が 10 倍になったとき, 期待度数の信頼性は以前より増えており, ここからのズレはより統計的に意味を持つことになる.

多項分布の検定の非筒の例として

$$H_0 : p_i = p_i^*,\ i = 1, 2, \ldots, k$$

の様に帰無仮説が単純仮説の場合, つまり特定の $p_i^*, i = 1, 2, \ldots, k$ が想定されている場合を考える．この場合は, 期待度数は $\hat{x}_i = np_i^*$ となるので, 検定等計量

$$W = \sum_{i=1}^{k} \frac{(x_i - np_i^*}{np_i^*}$$

を自由度 $k - 1$ のカイ二乗分布の上側 100 ％点と比較して, これより大きければ帰無仮説は棄却となる.

後半の仮説を考えてみよう．この場合, 仮説「20 代の客,　30 代の客, それ以外の年代の客を割合が, 2:1:1 である」は,

$$H_0 : p_1 = 1/2,\ p_2 = 1/4,\ p_3 = 1/4$$

という単純仮説になる．期待度数はそれぞれ, $\hat{x}_1 = 97/2, \hat{x}_2 = 97/4, \hat{x}_3 = 97/4$ となるので, 統計量の値は

$$W = \frac{(52-97/2)^2}{97/2} + \frac{(30-97/4)^2}{97/4} + \frac{(15-97/4)^2}{97/4} \sim 5.144$$

となる．自由度は 2 であるので, 5 %検定するために, 自由度 2 のカイ二乗分布の上側 5 %点 5.991 と比較すると, これより小さいので帰無仮説は採択される．

$\boxed{\textbf{ex.}}$ 遺伝の問題, メンデルの法則

ある花の品種改良の実験で標本数 $n = 217$ の株を得た．

1.	緑の斑点のあるマゼンダの花	120 株
2.	赤の斑点のあるマゼンダの花	48 株
3.	緑の斑点のあるマーガレットの花	36 株
4.	赤の斑点のあるマーガレットの花	13 株

メンデルの法則によると, これらの花は 9 : 3 : 3 : 1 の割合で得られる．この実験結果はメンデルの法則に適合するか？

モデル: ある変数 X について n 個の標本がとられ k 個の項に分類されているとする．($k = 2$ の時 2 項分布)

各項に入る確率が $p_1, p_2, \ldots, p_k$ ならば $p_1 + p_2 + \cdots + p_k = 1$　(2 項分布のときは $p_1 + p_2 = 1$)
n 個の標本が各項に入る期待値は

$$e_1 = np_1, e_2 = np_2, \ldots, e_k = np_k$$
(2 項分布のときは $e_1 = np_1, e_2 = np_2$)

実際の実験結果　$O_1, O_2, \ldots O_k$

仮説 H_0	:	変数 X は各項に入る確率が $p_1, p_2, \ldots, p_k$ という確率分布に従う．
対立仮説 H_1	:	変数 X は, H_0 の確率分布に従わない．

項	1	2	3	$\ldots$	k	計
確率 p_i	p_1	p_2	p_3	$\ldots$	p_k	1
期待値 e_i	$e_1 = np_1$	$e_2 = np_2$	$e_3 = np_3$	$\ldots$	$e_k = np_k$	n
観測度数 O_i	O_1	O_2	O_3	$\ldots$	O_k	n

χ^2(カイ 2 乗) の値を計算する．

$$(1) \qquad \chi^2 = \sum_{i=1}^{k} \frac{(e_i - O_i)^2}{e_i}$$

実験結果が適合すると χ^2 の値は小さい．χ の値がある値 χ_0^2 より大きいと H_0 は棄却される．

$$\text{棄却域} \quad : \quad \chi^2 > \chi_0^2$$

- $\alpha = 0.05$ で適合検定する．

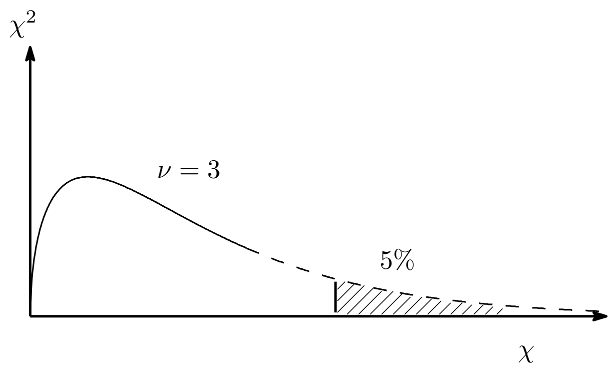
  表) χ^2 分布より自由度 $\nu = k - 1$ の値 χ_ν^2 を見つける．

【例の解法】

項	1	2	3	4	計
確率 p_i	9/16	3/16	3/16	1/16	1
期待値 e_i	122	41	41	14	217
観測度数 O_i	120	48	36	13	217

$\alpha = 0.05$ で適合検定する.
自由度 $\nu = k - 1 = 3$ より $\chi_0^2 = 7.81$. (χ^2 分布表を見る.)

$$\chi^2 = \frac{(122-120)^2}{122} + \frac{(41-48)^2}{41} + \frac{(41-48)^2}{41} = 1.9$$



χ^2 の値は χ_0^2 よりはるかに小さいのでこの実験はメンデルの法則に適合する.

2.7.2 独立性の検定

$\boxed{\textbf{ex.}}$ 学生の成績と外見は関係するのか？

120 人の学生を外見と成績の 2 つの変数によって分類した. 外見では「魅力的」,「普通」,「魅力的でない」の 3 つのカテゴリー, 成績では「高い」,「やや高い」,「やや低い」,「低い」の 4 つのカテゴリーで分けた.

Y	X				計
	高い	やや高い	やや低い	低い	
魅力的	14	11	10	5	40
普通	10	16	16	14	56
魅力的でない	3	4	7	10	24
計	27	31	33	29	120

- 独立性の検定をする.

$$\begin{aligned} &\text{仮説 } H_0 &:&\ X \text{ と } Y \text{ は独立 (外見と成績は関係ない)} \\ &\text{対立仮説 } H_1 &:&\ X \text{ と } Y \text{ は独立でない (関係あり)} \end{aligned}$$

- 解法

仮説 H_0 の下で X と Y は独立であるから

Y	X				計
	高い	やや高い	やや低い	低い	
魅力的	9	10	11	10	40
普通	13	14	15	14	56
魅力的でない	5	6	7	6	24
計	26	30	33	30	120

一行目

$$\begin{cases} 120 \times \dfrac{27}{120} \times \dfrac{40}{120} & = & 40 \times \dfrac{27}{120} & = & 9 \\ & & 40 \times \dfrac{31}{120} & = & 10 \\ & & 40 \times \dfrac{33}{120} & = & 11 \\ & & 40 \times \dfrac{29}{120} & = & 10 \end{cases}$$

2 行目, 3 行目も同様に計算する.

- $\alpha = 0.05$ で独立性を検定する.

$$\begin{aligned} \text{自由度} &= (\text{行の数-1}) \times (\text{列の数-1}) \\ &= 3 \times 2 \\ &= 6 \end{aligned}$$

このとき $\chi_0^2 = 12.59$.

$$\begin{aligned} \chi^2 = & \frac{(9-14)^2}{9} + \frac{(10-11)^2}{10} + \frac{(11-10)^2}{11} + \frac{(10-5)^2}{10} \\ & + \frac{(13-10)^2}{13} + \frac{(14-16)^2}{14} + \frac{(15-16)^2}{15} + \frac{(14-14)^2}{14} \\ & + \frac{(5-3)^2}{5} + \frac{(6-4)^2}{6} + \frac{(7-7)^2}{7} + \frac{(6-10)^2}{6} \\ & \fallingdotseq 10.7 < 12.89 \end{aligned}$$

H_0 は採択. 従って学生の外見と成績は関係あるとはいえない.

2.8 検定の演習問題

(i) x は未知の平均 μ, 標準偏差 $\sigma = 6$ の正規分布に従うとする. 大きさ 16 の標本をとり, $\overline{x} = 33$ を得たとき,

(a) 仮説 $H_0 : \mu = 30$ を, 対立仮説 $H_1 : \mu > 30$ に対して検定せよ.

(b) しばらくしてから, 大きさ 32 の第 2 の標本をとり, $\overline{x} = 29$ を得た. 母集団平均 μ がこの間に変化したとみなせるか. 平均は変化していないという仮説を, 平均は小さくなったという対立仮説に対して検定せよ.

(ii) ある市で 200 人の自動車所有者の標本を調べたら, そのうち 48 人が期限切れの運転免許証を持っていた. 前年度の期限切れ運転免許証所持率は 0.30 であった. これらのデータを用いて, 次の問題を解け.

(a) 母集団割合は $p = 0.30$ であるとの仮説を検定せよ.

(b) 母集団割合は $p = 0.30$ であるとして, 大きさ 50 の標本を毎日とる場合の $\widehat{p}$ に対する管理図の管理限界を求めよ.

(iii) 1974 年 4 月 27 日付の Newsweek に報道された実験は, マリファナが性活動に及ぼす影響を調べたもので, セントルイスの生殖研究協会によって行われた. 実験には 20 人の健康な若者が選ばれ, 1 週間に少なくとも 4 日, 最低 6 週間にわたりマリファナを吸わせたが, この期間中他の薬は一切使用させなかった. 対照群として, マリファナを吸わない 20 人の若者が比較のために選ばれた. 使用した性活動の尺度は, 血液中の男性ホルモン “テストステロン” の量であった. 添字の 1 と 2 はそれぞれマリファナ群と非マリファナ群に対応するとし, 実験の結果, 次のデータを得たとする.
$\overline{x}_1 = 416, \overline{x}_2 = 742, s_1 = 152, s_2 = 130.$

(a) 仮説 $H_0 : \mu_1 = \mu_2$ を検定せよ.

(b) 同じ仮説を小標本法によって検定せよ.

(iv) 1973 年 10 月 22 日付の Time の報道によると, 酒を飲めば顔が赤くなるという東洋人の主張が真実か否かを調べる研究がノースカロライナ州のチャペルヒルで行われた. この研究では適度な飲酒習慣をもつ東洋系アメリカ人 24 人と同じ程度の飲酒習慣をもつヨーロッパ系アメリカ人 24 人とを比較した. 結果は次のようであった.

東洋系アメリカ人は 24 人中 17 人が赤くなった.
一方, ヨーロッパ系アメリカ人は 24 人中 3 人が赤くなった.

仮説 $H_0 : p_1 = p_2$ を検定せよ. ここで p_1, p_2 は両グループにおける割合を表す.

(v) 以下の数値は $\mu = 3.0, \sigma = 0.5$ の正規母集団からの標本抽出によって得られたものである. おのおの 25 個の標本が 2 人の実験者 A, B によってとられた. これらのデータを用いて, 次の問題を解け.

(a) 全データを 1 組のデータとみなして, 仮説 $H_0 : \mu = 3.0$ を対立仮説 $H_1 : \mu \neq 3.0$ に対して検定せよ.

(b) これら 50 個の観測値から管理図を作れ.

(c) $\sigma_A = \sigma_B = \sigma = 0.5$ を用いて, $\mu_A = \mu_B$ なる仮説を検定せよ.

(d) 標本偏差が既知ではないとして (c) を解け.

(e) A のデータと t 分布を用いて, $H_0 : \mu_A = 3.3$ を $H_1 : \mu < 3.3$ に対して検定せよ.

(f) t 分布を用いて, 仮説 $H_0 : \mu_A = \mu_B$ を検定し, このときの t の値を (d) の z の値と比べよ.

A	3.5	3.4	2.8	2.6	2.3	2.4	2.9	2.9	3.1	3.3	3.9	3.0	2.9
B	2.9	3.6	3.1	2.5	2.6	3.6	2.9	2.1	2.6	2.5	2.3	2.6	3.3

3.0	2.5	2.7	3.9	2.6	2.8	3.2	3.6	2.7	3.0	3.8	3.1
3.9	2.7	2.4	2.8	3.0	3.3	3.7	3.3	3.1	2.7	2.4	3.3

(vi) 正しい硬貨を 2 回投げ, 2 回続けて表が出たら, ある仮説を棄却するとき, 2 種類の過誤の大きさは, それぞれいくらか?

(vii) 1 枚の硬貨を 2 回投げ, そのとき表の出た回数を x とする. 表に出る確率を p とし, x の値によって仮説の検定を行うものとして, 仮説 $H_0 : p = 0.5$ を対立仮説 $H_1 : p = 0.7$ に対して検定せよ. 検定の棄却域として $x = 2$ を選ぶとき, この検定での 2 種類の過誤の大きさをそれぞれ求めよ.

(viii) レンガの買い手は近頃のレンガの品質はどうも低下しているようだと思っていた. 過去の経験では, レンガの平均破壊強度は 400 ポンドで, 標準偏差は 20 ポンドであった. 100 個のレンガの標本をとってテストしたら, その平均は 390 ポンドであった. 平均品質は低下しているという対立仮説に対して, それは変わっていないという仮説を検定せよ.

(ix) 生物学者がある種の昆虫の 50% が殺せるように殺虫液を調合した. この液をこの種の昆虫 200 匹に吹きかけたところ, 120 匹が死んだ. 液の調合はうまくいったと結論してよいか.

(x) 50 人ずつの小学校の生徒 2 組に 2 通りの違った方法で読み方を教えた. 授業後読み方の試験を行い, 次の結果を得た. $\overline{x}_1 = 73.4, \overline{x}_2 = 70.2, s_1 = 9, s_2 = 10$. この結果から仮説: $\mu_1 = \mu_2$ を検定せよ.

(xi) ある都市におけるテレビの視聴率調査で, ある番組を男子は 200 人中 56 人が好まなかったのに対し, 女子は 300 人中 75 人がそれを好まなかった. 両者の嗜好に実際上差があるといえるだろうか.

参考文献

[1] 初等統計学：ポール G. ホーエル 著, 浅井 晃, 村上 正康 訳, 培風館.

[2] 統計解析入門：篠崎信雄 著, サイエンス社.

[3] やさしい統計学入門: 田栗正章, 藤越康祝, 柳井晴夫, C.R. ラオ 著, 講談社.

[4] やさしい統計: 吉原健一, 金川秀也 著, 培風館.

[5] 統計学入門：基礎統計学☒, 東京大学出版会.

[6] 統計学: 日本統計学会編, 東京図書.